inference. Also, while implementing the Lauritzen-Spiegelhalter algorithm, we developed an efficient method for finding maximal cliques in a triangulated graph, a necessary step in the application of the algorithm.

The Lauritzen-Spiegelhalter algorithm, in addition to the Pearl algorithm, which we implemented earlier, has been incorporated into KNET, a general knowledge engineering shell for constructing decision-theoretic medical expert systems [2]. We have developed a set of over 100 randomly generated bench-mark belief networks for formal evaluation of the Pearl and Lauritzen-Spiegelhalter algorithms, and comparison with other exact and approximation algorithms. We have commenced formal evaluation of the Pearl and Lauritzen-Spiegelhalter inference algorithms; preliminary results have given us important insights into the network topologies in which each of the algorithms performs more efficiently [12]. We are in the process of developing and implementing a method that allows us to use both algorithms simultaneously in performing belief network inference. We anticipate that the resulting synthetic algorithm will perform more efficiently than either algorithm alone.

We have designed a temporal probabilistic representation and inference method [25]. We have developed a prototype knowledge base and system that reasons about the causes of temporally-qualified symptoms in a patient. Reasoning about time can considerably increase the computational time complexity of probabilistic inference and our current techniques are likely to be inadequate. Thus, we currently are attempting to develop more efficient temporal inference methods.

Special-Case Algorithms

During the past year we have explored belief-network precompilation as a special-case method [27]. Precompiling a belief network consists of performing probabilistic inference for cases that are likely to be encountered and then storing the results, indexed by the evidence of the case. When the system is given a patient case, it first quickly checks to see it is has been stored. If so, it returns the probability distribution over the potential diseases almost immediately. If not, it solves the case more slowly with a belief-network inference algorithm. Work to date has demonstrated the feasibility of the general precompilation concept and the feasibility of caching incomplete evidence sets (that is, sets of evidence for which some values are unknown), or evidence sets selected by their relative expected utilities instead of expected joint probabilities. A theoretical justification of this method's success has been derived, and will be expanded.

Research work in the QMR-DT (Quick Medical Reference, Decision-Theoretic) project has explored probabilistic inference on a large medical knowledge base. During the past year, we developed a probabilistic version of the QMR decision-support system for diagnosis in internal medicine. Knowledge of over 600 diseases and 4000 manifestations in the QMR knowledge base is incorporated into the probabilistic interpretation of the knowledge base used by QMR-DT. In light of the size of the QMR-DT knowledge base, our approach to developing a pragmatic diagnostic system is first to make assumptions to reduce the inferential and representational complexity. We are now in the process of evaluating and incrementally modifying these assumptions, based on the performance of the system against QMR and patient cases with known diagnoses.

We have developed an algorithm that can perform probabilistic inference in QMR-DT in $O(n\, m^- 2^{m^+})$, where n is the number of diseases (n is approximately 600), m^- is the number of negative findings, and m^+ is the number of positive findings [19]. This development is a significant improvement over the straightforward inference algorithm, which is $O(2^n)$. In practice, the number of positive findings is far less than the number of diseases. The current algorithm, as implemented in LightSpeed Pascal on a Macintosh II, can score cases with 9 positive findings in less than one minute.

Approximation Algorithms

As part of our work on the value of bounding, we have developed and explored the modulation of the completeness of probabilistic inference by decomposing a problem into a set of inference subproblems, and by ordering the solution of these problem components by their expected impact on the overall probabilistic inference task. Our method, called bounded cutset conditioning [28], is a graceful analogue to the method of conditioning for probabilistic entailment in belief networks as developed by Pearl. Preliminary analysis indicates that this method can increase inference speed by a factor of 2 or 3 in some belief networks.

We also have developed a randomized approximation scheme for belief network inference [13]. Given a set of evidential variables, the algorithm computes posterior probabilities, with high probability, to within a prespecified error. The method combines Monte Carlo techniques, area-estimation strategies, and convergence analysis for time-reversible Markov chains.

I.B.3 Probability assessment

KNET is a general-purpose environment for constructing probabilistic medical expert systems [2, 17, 24]. Such networks serve as graphical representations for probabilistic models. KNET differs from other tools for expert-system construction in that it combines a direct-manipulation visual interface with a formal probabilistic scheme for the management of uncertain information and inference. The KNET architecture defines a complete separation between the hypermedia user interface on the one hand, and the representation and management of expert opinion on the other. KNET has been fully implemented and debugged, and currently runs on Macintosh II hardware. The system offers a choice of algorithms, some approximate and others exact, for probabilistic inference. We have used KNET to build consultation systems for lymph-node pathology [20], bone-marrow transplantation therapy [23], clinical epidemiology [1], and alarm management in the intensive-care unit [12]. KNET imposes few restrictions on the interface design. Indeed, we have rapidly prototyped several direct-manipulation interfaces that use graphics, buttons, menus, text, and icons to organize the presentation of static and inferred knowledge.

In the past year, research on *similarity networks* [20] has advanced in three areas. First, the theory of similarity networks was formalized. In particular, necessary and sufficient conditions for consistency among *local* Bayesian belief networks in the similarity network were derived. In addition, it was proved that the ad hoc procedure for construction the *global* belief networks from the collection of local belief networks is sound in the sense that any assertion of conditional independence implied by the global belief network can be derived from the structure of the local belief networks [26]. Second, the similarity network representation was used to build a knowledge base for the *entire* domain of lymph node pathology [20]. We found that the use of similarity networks greatly simplified the assessment of dependencies among findings. Also, the time to assess probabilities was reduced by approximately a factor of ten [26]. Third, the knowledge base was evaluated using a decision-theoretic metric [6]. We found that the diagnostic accuracy of the new knowledge base was significantly better than the accuracy of the knowledge base constructed without the use of similarity networks. In addition, we found that the new knowledge base performed as well as the expert within the noise level of the experiment .

I.C. Publications

<u>Articles published:</u>

1. R. M. Chavez and H. P. Lehmann, REFEREE: A belief network that helps evaluate the credibility of a randomized clinical trial, In: *Proceedings of the 1988 AAAI Workshop on Artificial Intelligence in Medicine*, Stanford, California, 1988.
2. R. M. Chavez and G. F. Cooper, KNET: Integrating hypermedia and Bayesian modeling, In: *Proceedings of the AAAI Workshop on Uncertainty in Artificial Intelligence*, Minneapolis, Minnesota, August 1988.
3. G.F. Cooper, Computer-based medical diagnosis using belief networks and bounded probabilities, In: P. Miller (Ed.), *Readings in Medical Artificial Intelligence*, (Springer-Verlag, New York, 1988).
4. G.F. Cooper, A method for using belief networks as influence diagrams, *Proceedings of the AAAI Workshop on Uncertainty in Artificial Intelligence*, Minneapolis, Minnesota, August 1988.
5. G.F. Cooper, invited commentary on: Lauritzen, S. and Spiegelhalter, D., Local computations with probabilities on graphical structures and their application to expert systems, *Journal of the Royal Statistical Society*, B, 50, 1988.

6. D.E. Heckerman, An empirical comparison of three scoring schemes, *Proceedings of the AAAI Workshop on Uncertainty in Artificial Intelligence*, Minneapolis, Minnesota, August 1988.

7. E.J. Horvitz, Reasoning under varying and uncertain resource limitations, In: *Proceedings of American Association for Artificial Intelligence*, Minneapolis, MN, August, 1988.

8. E.J. Horvitz, J.S. Breese, and M. Henrion, Decision theory in expert systems and artificial intelligence, *Journal of Approximate Reasoning*, 2: 247-302, 1988.

9. Lehmann, H., Knowledge acquisition for probability-based expert systems, The Symposium on Computer Applications in Medicine, November 1988.

10. H.J. Suermondt, and G.F. Cooper, Updating probabilities in multiply connected belief networks, In: *Proceedings of the AAAI Workshop on Uncertainty in Artificial Intelligence,* Minneapolis, Minnesota, August 1988.

Articles to appear:

11. T. Barsalou, R. M. Chavez, and G. Wiederhold, Hypertext interfaces for decision-support systems -- a case study, *MEDINFO-89*, November, 1989.

12. I. A. Beinlich, H. J. Suermondt, R. M. Chavez, and G. F. Cooper, The ALARM monitoring system: A case study with two probabilistic inference techniques for belief networks, *Proceedings of the Artificial Intelligence in Medicine Conference*, June,1989.

13. R. M. Chavez and G.F. Cooper, A fully polynomial randomized approximation scheme for the Bayesian Inferencing Problem, *Networks*, 1989.

14. G.F. Cooper, Expert systems based on belief networks -- Current research directions, *Applied Statistical Models and Data Analysis*.

15. G.F. Cooper, The computational complexity of probabilistic inference using belief networks, *Artificial Intelligence.*

16. E.J. Horvitz, G.F. Cooper, and D.E. Heckerman, Reflection and action under scarce resources: Theoretical principles and empirical study, *Proceedings of the International Joint Conference on Artificial Intelligence*, July, 1989.

Articles recently submitted for publication:

17. R. M. Chavez, Hypermedia and randomized algorithms for medical expert systems, The Symposium on Computer Applications in Medicine, November 1989.

18. R. M. Chavez, An empirical evaluation of a randomized algorithm for probabilistic inference, The AAAI Workshop on Uncertainty in Artificial Intelligence, July 1989.
19. D.E. Heckerman, A tractable inference algorithm for diagnosing multiple diseases, The AAAI Workshop on Uncertainty in Artificial Intelligence, July, 1989.
20. D.E. Heckerman, E.J. Horvitz, B. Nathwani, Toward effective normative decision systems: Update on the Pathfinder project, The Symposium on Computer Applications in Medicine, November 1989.
21. E.J. Horvitz, D.E. Heckerman, K.C. Ng, B. Nathwani, Heuristic abstraction in the decision-theoretic Pathfinder system, The Symposium on Computer Applications in Medicine, November 1989.
22. H.J. Suermondt and G.F. Cooper, Probabilistic inference in multiply connected belief networks using loop cutsets, *International Journal for Approximate Reasoning.*
23. H.J. Suermondt and M.D. Amylon, Probabilistic prediction of the outcome of bone-marrow transplantation, The Symposium on Computer Applications in Medicine, November 1989.

Additional research reports:

24. R. M. Chavez, Hypermedia and randomized algorithms for probabilistic expert systems, Ph.D. thesis proposal, March,1989.
25. G.F. Cooper, E.J. Horvitz, and D.E. Heckerman, A Model for Temporal Probabilistic Reasoning, Report KSL-88-***, Knowledge Systems Laboratory, Stanford University, July, 1988.
26. D.E. Heckerman, Probabilistic Similarity Networks, Ph.D. dissertation, Medical Information Sciences, Stanford University (anticipated date of completion: June, 1989).
27. E.H. Herskovits and G.F. Cooper, Algorithms for belief network precomputation, Report KSL-89-35, Knowledge Systems Laboratory, Stanford University, April, 1989.
28. E.J. Horvitz, H.J. Suermondt, G.F. Cooper, Bounded cutset conditioning: An incremental-refinement approach to inference under uncertain resources, Report KSL-88-36, Knowledge Systems Laboratory, Stanford University, December, 1988.

Presentations:

1. Chavez, A fully-polynomial randomized algorithm for probabilistic inference, An invited colloquium talk at the Center for Research on Computing Technology, Division of Applied Sciences, Harvard University, Cambridge, MA, February 1989.

2. Cooper, Addressing the computational complexity of medical diagnosis, An invited colloquium, talk at the Computer Science Department, Washington University, St. Louis, MO, December,1988.
3. Horvitz, Some fundamental problems and opportunities from the standpoint of rational agency," AAAI Spring Symposium on Artificial Intelligence and Rational Agency, Stanford University, March, 1989.
4. Horvitz, Inference under scarce resources, An invited colloquium, talk at the Computer Science Department, Washington University, St. Louis, MO, November,1988.

I.D. Funding Support

NSF Grant IRI-8703710
Pragmatic Approaches to Reasoning Under Uncertainty
1987-1990

U.S. Army Research Office Grant 25514-EL
Computational Techniques for Probabilistic Inference
1988-1991

II. Collaboration

During the past year we have continued to maintain contact with a number of researchers who are interested in the research goals described in Section I. These include Prof. Ross Shachter (Engineering-Economic Systems at Stanford), Prof. Harry Lewis (Division of Applied Sciences, Harvard University), Dr. Randolph Miller (Medical Informatics, University of Pittsburgh), Profs. Peter Szolovits and Ramesh Patil (Clinical Decision Making Group, M.I.T.), Prof. Max Henrion (Carnegie-Mellon University), Dr. David Spiegelhalter (Medical Research Council, England), and Dr. Stig Andersen (Aalborg University, Denmark). The SUMEX-AIM computer resource has greatly facilitated our maintaining communication with these researchers during the year.

The majority of our group participated in the AI in Medicine Spring Symposium held last year.

III. Research plans

III.A. Reasoning about inference tradeoffs

So far, our work on the value of probabilistic inference has focused on the value of computation in the setting of probability bounding algorithms. Theoretical work in the coming year will explore the possibility of tractable value-of-information calculations and approximations for several families of probability distributions. Also, in the coming year, we will work to extend the bounded-conditioning approach to inference under uncertain resource constraints to handle multiple pieces of evidence. We will seek to make the algorithm more robust by combining the method with a stochastic simulation approach for determining weights on alternative instances. We also plan to construct an inference simulation system for gathering information about the convergence of inference algorithms under different evidential settings. The goal of the off-line simulation is to collect parameters that can be used in real-time value of computation calculations for controlling inference. Finally, we plan to evaluate the decision-theoretic control of the construction of relatively small decision models from a large homogeneous belief network, in the internal medicine (QMR-DT knowledge base) setting.

III.B. Efficient probability inference algorithms

We plan to continue the formal evaluations of the algorithms that we have implemented, and submit the results for publication. We intend to apply the results of these evaluations in our work on combining algorithms for inference in a given belief network, and on developing new algorithms.

We will extend our current work on temporal belief-network representations and attempt to design more efficient methods for temporal probabilistic inference.

Other work will include refining and designing algorithms for exact and approximate inference on the QMR-DT belief-network. In particular, research in this area will focus on testing the validity of assumptions in the QMR-DT knowledge base and relaxing the sensitive assumptions by augmenting the inference algorithms or the knowledge base.

III.C. Probability assessment and knowledge acquisition

The QMR-DT knowledge base currently consists of a two-level belief network in which diseases are said to cause findings. We plan to augment the QMR-DT knowledge base to more realistically model the causal structure linking findings, diseases, and intermediate pathophysiological states.

During the summer of 1989 we will complete our work on similarity networks, which are used to increase the efficiency of acquiring probabilities directly from experts. We will turn our efforts toward developing a system that uses a medical database and a set of initial constraints to automatically construct a belief network.

IV.A.7. VentPlan Project

VentPlan Project: Combining Quantitative and Qualitative Techniques for Data Interpretation and Therapy Planning in the Intensive Care Unit

Lawrence Fagan, M.D., Ph.D.
Department of Medicine
Stanford University

Adam Seiver, M.D.
Department of Surgery
Veterans Administration Hospital
Palo Alto, California

Lewis Sheiner, M.D.
Department of Laboratory Medicine
University of California, San Francisco

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

We are designing a data-interpretation and therapy-planning system for the intensive care unit (ICU). Fundamental research issues in temporal reasoning are associated with the ICU application area including assimilation of incoming data, representation of time-oriented intervals, and description of ongoing physiological processes [Fagan 84]. In addition, in ICUs of the 1990's, many more physiological measurements will need to be collected at frequent intervals, and increased access to the current medical record in coded format will be possible. The goal of our system is to gather multiple measurements as they become available in the ICU, determine the meaning of the measurements with reference to the particular clinical context and the patient's individual response patterns, and to suggest alternative settings for the mechanical ventilator that supports the patient's respiratory function following surgery.

Our approach is to coalesce quantitative and qualitative models. Researchers have built elaborate mathematical models of the respiratory and cardiac systems, but it is impractical to base the entire reasoning process on complex equations. These models may take too long to process data or make assumptions about the clinical situation that are unwarranted. There exist many ways to represent the static relationships between different measurements and their related diagnoses, but those systems cannot easily represent complex temporal information. We wish to combine the qualitative techniques in order to build a system that utilizes the advantages of each type of model.

Adapting our approach for planning treatments for cancer patients with unusual clinical courses [Langlotz 87], we will use strategic knowledge to create patient-specific specializations of standard treatment plans. These plans will be used as general guidelines by the mathematical model to start a search for optimal treatment interventions. In a similar manner, a Bayesian net will be used to specify the initial ranges for parameters used by the mathematical model as it attempts to adapt to the the incoming data. The Bayesian net can also be used to provide a type of "smart alarms" where the interpretation of data is influenced by the specific diagnoses attributed to the patient. We will use decision analytic methods to evaluate and explain the various treatment options available at any point in time. The long-term goal of this project is to embed the decision-making components within the data management tasks of the ICU.

B. Medical Relevance and Collaboration

The problem of too much data being generated in the ICU is well recognized. Originally, monitors were designed to provide more objective assessments of the physiology of the patients in life-threatening situations. However, as more and more measurements became available, the ability of clinicians to assimilate the data began to drop. Expert systems can be designed to sort through the data, recognize untoward events in context and help with therapy selection. An early version of this was the VM system, which was based on extensions to the production rule framework. The current research has far broader goals, including real-time response using multiple methods for reasoning and the use of integrated mathematical models. This has led to a three way collaboration between the VA hospital which is installing a data management system for a new Surgical ICU (Seiver), the Medical Computer Science Group where investigations in qualitative-quantitative reasoning is taking place (Fagan), and the mathematical modeling group at U.C.S.F (Sheiner). In addition, we are continuing informal discussions with Barbara Hayes-Roth's group (see description of Guardian project).

C. Highlights of Research Progress

We have built a prototype version of the system which is being tested on realistic data derived from ICU measurements. The long range therapy planning components have not yet been created, but we have linked the mathematical models with the Bayesian network and the decision-theoretic component. The entire system is held together with a control algorithm that selects the order that modules show be invoked.

The central design principle of the VentPlan project is to develop a method for combining qualitative and quantitative modeling techniques. *Mathematical models* require the relationships between variables to be expressed as equations. Describing how a particular disease state—such as new onset of pulmonary embolus—affects a variety of respiratory and cardiac parameters is difficult. We are using *probabilistic causal networks* (also known as Bayesian networks) to represent qualitative relationships of the form: if A is present and B is high, then C is low with some probability. A complex set of relationships of this type can be built up into a network.

The mathematical model is able to estimate model parameters using patient data in order to generate patient-specific predictions of model variables—for example, partial pressure of CO2 in arterial blood (paCO2). The patient-specific mathematical model can be used to predict future measurements, which is difficult to do using the network formalism. Two other components complete the system. The first component is the *control algorithm*, which examines the incoming data and determines which of the modules should process the information. The *utility model* is used to encode physicians' preferences for treatment goals—for example, representing the tradeoff between providing a sufficiently high level of oxygen to the tissues and causing toxicity from having too high a level of inspired oxygen. VentPlan uses the utility model in the therapy planning process to rank plans that are created by the mathematical model.

VentPlan Control Algorithm. The control algorithm is activated each time that new data become available. Each iteration begins with the acquisition of new data, such as a blood-pressure reading. If the data do not directly correspond to quantities in the physiologic model, the belief network is evaluated. The network takes these quantitative inputs and the qualitative information—for example, the diagnosis—and derives prior-probability distributions for the shared model parameters. The mathematical model updates the parameters by combining the probability distributions with quantitative observations—for example, measurements of oxygen concentration in the blood. The mathematical model is then used to search for an optimal plan. The program simulates therapy plans under consideration by running the model, then ranks them using the plan evaluator. The therapy plans are sequentially modified to find the best plan. The calculated optimal plan is the recommended plan, which is compared to the current plan; this comparison is presented to the user.

VentPlan continuously monitors the stream of data measurements, reapplying this control sequence to refine the model and to recalculate the optimal plan as new information becomes available. New data are compared with the expected values derived from the mathematical model or the belief network; in this way, the program accomplishes expectation-driven analysis of the incoming data.

Belief Nets. Medical knowledge is represented in a belief network as a directed, acyclic graph in which nodes represent domain variables and arcs show important dependencies among these variables . Probabilities are attached to nodes and to groups of arcs. Prior probabilities are considered as static facts in this knowledge base. The conditional probabilities for a node are equivalent to rules of the form: "IF (parent node values) THEN (node values), with a stated probability."

Typical inputs are "mild congestive heart failure," "normal respiratory volume," and "decreased temperature." The belief network module provides probability distributions for the parameters that it shares with the mathematical model based on the information computed by the network. The corresponding outputs are estimates (mean and standard deviation) for each of the parameters, for example, cardiac output 4.4 +/- 2.3 l/min, and oxygen consumption 180 +/- 120 ml/min.

Diagnosis nodes are at the top of the network. These nodes have no predecessors, and we assume they are mutually independent a priori. All diagnosis nodes are associated with a set of mutually exclusive and exhaustive values representing the presence, absence, or severity of a particular disease. *Measurement nodes* represent any available quantitative information. All continuous variables are represented categorically, with sets of discrete intervals dividing the value range. Depending on the necessary level of detail, three to six categories are used for each node. *Parameter nodes* are inferred entities that cannot be measured directly. Each parameter node corresponds to one of the parameters shared with the mathematical model and with the plan evaluator.

The probabilities in a belief network can represent objective as well as subjective knowledge. The network for the VentPlan architecture contains statistical data on prior probabilities, objective conditional probabilities computed from physiologic equations, and subjective assessments. It is necessary to obtain conditional probabilities for the states of a node given all different states of the parent nodes. A probability editor lets the user browse through this multidimensional probability matrix. An equation editor generates conditional probabilities for any deterministic relationship.

The current VentPlan network consists of 81 nodes. Eight nodes represent diagnoses corresponding to typical problems in ventilator care (for example, airway obstruction, infection, and heart failure); 18 nodes represent input measurements. Information for both types of nodes is placed in associated *report nodes*. These report nodes are influenced by error nodes, which implement error functions . These error nodes allow representation and detection of conflicting evidence. For example, some evidence might suggest that the patient has a high oxygen concentration, whereas other evidence might suggest that the oxygen concentration is low. Both pieces of information can be incorporated; their error functions will regulate the effect on the rest of the network. If the information on the low oxygenation is from a highly erratic sensor, it will be discounted, while information from a more reliable blood-gas analysis will have a greater influence.

The inference engine is based on the Lauritzen–Spiegelhalter algorithm for local probability computations on graphical structures . It finds a set of probability distributions consistent with the available evidence and with the conditional probabilities in the network. The algorithm rearranges a network into a tree structure suitable for fast updates. Details on the algorithm can be found in Appendix F. We assume that the discrete probability distributions are approximated by the normal distribution. This assumption enables the network module to calculate initial parameter estimates as means and variances of the equivalent normal distribution.

Mathematical Model. The mathematical model of the VentPlan prototype consists of a quantitative model and a variety of numerical routines . The model is written as a set of linked differential equations that represent the mechanisms of oxygen transport through the body. Variables in the model equations capture the time-varying concentrations of oxygen as that gas flows through the circulatory system. The model has a number of highly nonlinear aspects, so it requires solution by iterative numerical techniques. The model makes a number of assumptions in order to maintain simplicity, as discussed in the section on levels of model detail.

Parameters are used in the mathematical model to represent the relevant underlying mechanisms of the patient. Patient parameters often are not directly observable, or are costly to measure. Once their values are known, the underlying patient state is largely characterized. For example, we use the parameter cardiac output to represent the volume of blood pumped by the heart per unit time. This parameter is difficult to measure directly, but the modeling procedures can infer its value from other observations.

Inputs to the model consist of initial parameter values, various patient measurements, the ventilator settings (treatment plan), and the times at which the observations were obtained. Solution of the model equations (simulation of the model) gives a prediction about the patient state resulting from these inputs. The model simulation is also used in the tasks of parameter updating and plan optimization. Time is included directly in the system equations (the derivatives are taken with respect to time). Patient observations are time-stamped to allow reasoning about different patient contexts and retrograde fitting of model parameters. Predictions may be generated for any desired future time. If predictions are requested for times in the distant future (with respect to the time constants of the model), the equilibrium form of the differential equations is solved.

Prior-probability distributions for the shared VentPlan parameters are provided by the belief network. When direct observations of the system are scarce, these values form a basis for model predictions. As more measurements become available, the prior-probability distributions become less important. This process of updating system parameters is known as *parameter estimation.* Parameter estimation converts the general mathematical model to a patient-specific model, the latter being able to generate patient-specific predictions . The optimization procedure uses the patient-specific model to find the best plan, given a specific patient state. It accomplishes this task by minimizing an objective (search) function defined by the plan evaluator, which proceeds by numerical iteration. The optimization routines are able to optimize the ventilator settings individually or as a group.

Plan Evaluator. The plan evaluator provides a relative ranking of plans and their predicted consequences. This ranking is complicated by conflicting objectives, which are typical of medical decisions. In ventilator management, for example, increasing the percentage of oxygen in the breathing mixture can improve the patient's oxygenation status, but high concentrations of oxygen have toxic effects. The optimal oxygen concentration represents a tradeoff between an improvement in oxygenation and an increase in oxygen toxicity. A multi-attribute value function is used to determine the optimal combination of objectives . The attributes of the value model are the proposed therapy plan (ventilator settings) and selected model predictions. A cost for each of the attributes is determined using a value function. These costs are weighted and summed to obtain the overall cost for a plan.

D. Relevant Publications

1) Fagan, L., Kunz, J., Feigenbaum, E, and Osborn, J. Adapting a rule-based system for a monitoring task, in *Rule Based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project*, B. Buchanan and E. Shortliffe (eds.). Reading, MA: Addison-Wesley Publishing Co., 1984.

2) Langlotz, C., Fagan, L., Tu, S., Sikic, B., and Shortliffe, E. A therapy planning architecture that combines decision theory and artificial intelligence techniques. *Computers and Biomedical Research* 20:279-303, 1987.

3) Rutledge. G., Thomsen, G., Beinlich, I., Farr, B. Kahn, M., Sheiner, L., and Fagan, L. VentPlan: An architecture for combining qualitative and quantitative computation. Report KSL-89-04, January 1989.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

As described above, this project is a three-way collaboration between the Departments of Medicine, and Department of surgery at the VA Hospital, and U.C.S.F. As, such we will need electronic mail and networking facilities. In addition, we imagine strong interactions with other projects around the world with similar research goals. We have already been contacted by research groups in Holland, Scotland, and Norway. In addition, similar research projects are underway at Yale, Berkeley, and Chicago. We expect that the networking facilities may allow us to share test cases, and possibly knowledge bases.

B. Sharing and Interaction with Other SUMEX-AIM Projects

The Yale project mentioned above is associated with Perry Miller's group. We also expect considerable interaction with the ONCOCIN and other parts of the Heuristic Programming Project at Stanford.

C. Critique of Resource Management

The SUMEX staff have been quite helpful in the support of the various machines that have been used in this project so far. We believe that the current efforts of the SUMEX staff are quite appropriate for our research needs.

III. RESEARCH PLANS

Our basic research agenda is described above. The basic research issues underlying this project will extend for several years, leading to an implementation in the Veterans Administration Hospital in Palo Alto. This research will continue to need help assistance with local area networking, file service, and inter-network mail. We will need support for communications support within a project that is spread out over three geographical sites.

The SUMEX staff has been quite useful in providing support in other configurations of mainframe and workstations networked together. We anticipate that support for our unique collaborative arrangement will be equally superb.

IV.B. National AIM Projects

The following group of projects is formally approved for access to the AIM aliquot of the SUMEX-AIM resource. Their access is based on review by the AIM Advisory Group and approval by the AIM Executive Committee.

IV.B.1. INTERNIST-I/QMR Project

J. D. Myers, M.D.
University Professor Emeritus (Medicine)

Randolph A. Miller, M.D.
Associate Professor of Medicine
Chief, Section of Medical Informatics

University of Pittsburgh
1291 Scaife Hall
Pittsburgh, Pa., 15261

I. SUMMARY OF RESEARCH PROGRAM

A. Project rationale

The principal objective of this project is the development of a high-level computer diagnostic program in the broad field of internal medicine as an aid in the solution of complex and complicated diagnostic problems. To be effective, the program must be capable of multiple diagnoses (related or independent) in a given patient.

A major achievement of this research undertaking has been the design of a program called INTERNIST-I, along with an extensive medical knowledge base. This program has been used over the past decade to analyze many hundreds of difficult diagnostic problems in the field of internal medicine. These problem cases have included cases published in medical journals (particularly Case Records of the Massachusetts General Hospital, in the New England Journal of Medicine), CPCs, and unusual problems of patients in our Medical Center. In most instances, but by no means all, INTERNIST-I has performed at the level of the skilled internist, but the experience has highlighted several areas for improvement.

B. Medical Relevance and Collaboration

The program inherently has direct and substantial medical relevance.

The development of the QUICK MEDICAL REFERENCE (QMR) under the leadership of Dr. Randolph A. Miller has allowed us to distribute the INTERNIST-I knowledge base in a modified format to over twenty other academic medical institutions. The knowledge base can thereby be used as an "electronic textbook" in medical education at all levels -- by medical students, residents and fellows, and faculty and staff physicians. This distribution is continuing to expand.

The INTERNIST-I program has been used in recent years to develop patient management problems for the American College of Physician's Medical Knowledge Self-assessment Program.

C. Highlights of Research Progress

C.1 Accomplishments this past year

For the record, it should be noted that grant support for the QMR project has come solely from the CAMDAT Foundation of Farmington, Conn., from the Department of Medicine of the University of Pittsburgh, and from Dr. Miller's NLM RCDA grant, NLM RO1 grant and NLM UMLS contract.

In 1987, the University of Pittsburgh was named recipient of a National Library of Medicine Medical Informatics Training Grant Award.

The group of us (Myers, Miller and Masarie) together with assigned residents in internal medicine and fellows in medical informatics are continuing to expand the knowledge base and to incorporate the diagnostic consultative program into QMR. The computer program for the interrogative part of the diagnostic program is the main remaining task. An editor for the QMR knowledge base, as modified from the INTERNIST-I knowledge base, has been written from scratch in Turbo Pascal by Dr. Masarie. The entire QMR program can be accommodated in, maintained (particularly edited) and operated on individual IBM PC-AT computers.

Our group has incorporated into the QMR diagnostic consultant program modifications and embellishments of the INTERNIST-I knowledge base, and will continue to do so over the next year by adding "facets" of diseases or syndromes. This addition and modification is expected to improve the performance of the diagnostic consultant program.

The medical knowledge base has continued to grow both in the incorporation of new diseases and the modification of diseases already profiled so as to include recent advances in medical knowledge. Several dozen new diseases have been profiled during the past year. The current number of diseases in the QMR knowledge base is 597, and 4260 possible patient findings are included.

C.2 Research in progress

There are four major components to the continuation of this research project:

1) The enlargement, continued updating, refinement and testing of the extensive medical knowledge base required for the operation of INTERNIST-I and the QMR modification.
2) Institution of field trials of QMR on the clinical services in internal medicine at the Health Center of the University of Pittsburgh. This has been accomplished in a limited fashion, which began in 1987; a "computer-based diagnostic consultation service" has been made available to attending physicians and house staff on the medical services of our two main teaching hospitals. Institutional Review Board (IRB) approval was granted to the service before it was initiated.

3) Expansion of the clinical field trials to other university health centers which have expressed interest in working with the system.
4) Adaptation of the diagnostic program and data base of INTERNIST-I and the QMR modification to subserve educational purposes and the evaluation of clinical performance and competence.

Current activity is devoted mainly to the first two of these, namely, the continued development of the medical knowledge base, and the implementation of the improved diagnostic consulting program, and preliminary evaluation of the diagnostic consultation service.

D. List of relevant publications

1) Bankowitz RA, McNeil MA, Challinor SM, Parker RC, Kapoor WN, Miller RA. A Computer-Assisted Medical Diagnostic Consultation Service: Implementation and Prospective Evaluation of a Prototype. Annals of Int Med. 1989, 110(10):824-832.
2) Bankowitz RA, McNeil MA, Challinor SM. Effect of a Computer-Assisted General Medicine Diagnostic Consultation Service on Housestaff Diagnostic Strategy. Proceedings of International Symposium on Medical Informatics and Education, Victoria, B.C., May 15-19,1989, pp. 219-223.
3) Giuse NB, Giuse DA, Miller RA. Learning by Doing: A Case Study in Medical Informatics. Proceedings of the Sixth National Symposium on Computers in Medical Education. Omaha, Nebraska. March 28, 1989.
4) Giuse NB, Giuse, DA, Miller RA. Computer assisted multi-center creation of medical knowledge bases. Proceedings of the 12th Annual Symposium on Computer Applications in Medical Care. IEEE Press. pp. 583-90, November 1988.
5) Giuse NB, Giuse DA, Miller RA. Medical knowledge base construction as a means of introducing medical students to medical informatics. Proceedings of the International Symposium on Medical Informatics in Education. Victoria, B.C., May 15-19, 1989, pp. 228-232.
6) Miller RA, et al. Preparing a Research Grant Proposal in Medical Informatics. Comp Biomed Res. 22(1):92-101, 1989.
7) Miller RA. Legal Issues Related to the Use of Medical Decision Support Systems. International J Clin Monitoring and Computing. 1989. (in press).
8) Berner ES, Brooks CM, Miller RA, Masarie FE Jr, Jackson JR. Evaluation Issues in the Development of Medical Decision Support Software. Evaluation and the Health Professions. Forthcoming 1990.
9) Lincoln M, Turner C, Hesse B, Miller RA. A Comparison of Clustered Knowledge Structures in Iliad and in Quick Medical Reference. Proceedings of 12th Annual Symposium on Computer Applications in Medical Care. IEEE Press, pp. 131-135, November 1988.

10) Miller RA, Masarie FE Jr. Use of the Quick Medical Reference (QMR) (R) Program as a Tool for Medical Education. Proceedings of the International Symposium on Medical Informatics and Education. Victoria, B.C. May 15-19,1989, pp. 247-252.

11) Parker RC, Miller RA. Creation of a Knowledge Base Adequate for Simulating Patient Cases: Adding Deep Knowledge to the INTERNIST-1/QMR Knowledge Base. Proceedings the International Symposium on Medical Informatics and Education. Victoria, B.C. May 15-19, 1989, pp. 281-286.

12) McNeil M, Parker R, Bankowitz R, Challinor S. Correlates of internal medicine as a residency choice among students at the University of Pittsburgh. Society of General Internal Medicine Annual Meeting, 1989. (abstract)

13) Parker, S, Kroboth F, Parker R, Hanusa B, Kapoor W. Development of an easily completed and scored physician satisfaction questionnaire for use by both inpatients and outpatients. Society of General Internal Medicine Annual Meeting, 1989. (abstract)

E. Funding support

1) Diagnostic-Internist: A Computerized Medical Consultant
   Randolph A. Miller, M.D.
   Associate Professor of Medicine
   Chief, Section of Medical Informatics
   University of Pittsburgh Department of Medicine
   National Library of Medicine - Development Award Research Career
   National Institutes of Health
   5 KO4 LM00084
   09/30/85 - 09/29/86 - $55,296
   09/30/86 - 09/29/87 - $55,296
   09/30/87 - 09/29/88 - $54,648
   09/30/88 - 09/29/89 - $54,864
   Support recommended for 1 additional year ending 09/29/90. The Amount to be determined annually.

2) Developing INTERNIST-I Knowledge Base into a Resource
   Randolph A. Miller, M.D.
   Associate Professor of Medicine
   Chief, Section of Medical Informatics
   University of Pittsburgh Department of Medicine
   National Library of Medicine
   National Institutes of Health
   1 RO1 LM04622
   09/30/87 through 09/29/90
   09/30/87 - 09/29/88 - $71,892
   09/30/88 - 09/29/89 - $99,639
   09/30/89 - 09/29/90 - $112,580

3) Pittsburgh Medical Information Sciences Training Program
Randolph A. Miller, M.D.
Associate Professor of Medicine
Chief, Section of Medical Informatics
University of Pittsburgh Department of Medicine
National Library of Medicine
National Institute of Health
5 T15 LM07059
07/01/87 through 06/30/92
07/01/87 - 06/30/88 - $153,454
07/01/88 - 06/30/89 - $221,511
07/01/89 - 06/30/90 - $265,930
07/01/90 - 06/30/91 - $278,092
07/01/91 - 06/30/92 - $276,596

4) Unified Medical Language System Support (UMLS)
Randolph A. Miller, M.D.
Associate Professor of Medicine
Chief, Section of Medical Informatics
University of Pittsburgh Department of Medicine
National Library of Medicine
N01-LM-8-3514
06/30/88 through 06/29/91
06/30/88 - 06/29/89 - $133,617
06/30/89 - 06/29/90 - $129,573
06/30/90 - 06/29/91 - $188,824

5) Research Proposal for Implementation of Chinese Version of QMR
Nunzia B. Giuse, M.D.
Research Associate
University of Pittsburgh Department of Medicine
Section of Medical Informatics
CAMDAT Foundation
06/01/88 - 12/31/89
$16,500

6) Camdat Support of Quick Medical Reference (QMR)
Nunzia B. Giuse, M.D.
Research Associate
University of Pittsburgh Department of Medicine
Section of Medical Informatics
CAMDAT Foundation
08/01/88 - 12/31/89
$21,300

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A,B. Medical Collaborations and Program Dissemination Via SUMEX

INTERNIST-I and QMR remain in a stage of research and particularly development. As noted above, we are continuing to develop better computer programs to operate the diagnostic system, and the knowledge base cannot be used very effectively for collaborative purposes until it has reached a critical stage of completion. These factors have stifled collaboration via SUMEX up to this point and will continue to do so for the next year or two. In the meanwhile, through the SUMEX community there continues to be an exchange of information and states of progress. Such interactions particularly take place at the annual AIM Workshop.

C. Critique of Resource Management

SUMEX has been an excellent resource for the development of INTERNIST-I. Our large program is handled efficiently, effectively and accurately. The staff at SUMEX have been uniformly supportive, cooperative, and innovative in connection with our project's needs.

III. RESEARCH PLANS

A. Project Goals and Plans

Continued effort to complete the medical knowledge base in internal medicine will be pursued including the incorporation of newly described diseases and new or altered medical information on "old" diseases. The latter two activities have proven to be more formidable than originally conceived.

B. Justification and Requirements for Continued SUMEX Use

Our use of SUMEX has declined with the adaptation of our programs to the IBM PC-AT. Nevertheless, the excellent facilities of SUMEX are expected to be used for certain developmental work. It is intended for the present to keep INTERNIST-1 at SUMEX for comparative use as QMR is developed here. We will not need the DEC 2060 beyond its anticipated phase-out in early 1989, but will require access to its replacement for mailing purposes and to maintain contact with the national medical informatics community.

C. Needs and Plans for Other Computing Resources Beyond SUMEX-AIM

Our predictable needs in this area will be met by our recently acquired personal work stations.

IV.B.2. MENTOR Project

MENTOR Project — Medical Evaluation of Therapeutic Orders

Stuart M. Speedie, Ph.D.
School of Pharmacy
University of Maryland

Terrence F. Blaschke, M.D.
Department of Medicine
Division of Clinical Pharmacology
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The goal of the MENTOR (Medical EvaluatioN of Therapeutic ORders) project is to design and develop an expert system for monitoring drug therapy for hospitalized patients that will provide appropriate advice to physicians concerning the existence and management of adverse drug reactions. The computer as a record-keeping device is becoming increasingly common in hospital-based health care, but much of its potential remains unrealized. Furthermore, this information is provided to the physician in the form of raw data which is often difficult to interpret. The wealth of raw data may effectively hide important information about the patient from the physician. This is particularly true with respect to adverse reactions to drugs which can only be detected by simultaneous examinations of several different types of data including drug data, laboratory tests and clinical signs.

In order to detect and appropriately manage adverse drug reactions, sophisticated medical knowledge and problem solving is required. Expert systems offer the possibility of embedding this expertise in a computer system. Such a system could automatically gather the appropriate information from existing record-keeping systems and continually monitor for the occurrence of adverse drug reactions. Based on a knowledge base of relevant data, it could analyze incoming data and inform physicians when adverse reactions are likely to occur or when they have occurred. The MENTOR project is an attempt to explore the problems associated with the development and implementation of such a system and to implement a prototype of a drug monitoring system in a hospital setting.

B. Medical Relevance and Collaboration

A number of independent studies have confirmed that the incidence of adverse reactions to drugs in hospitalized patients is significant and that they are for the most part preventable. Moreover, such statistics do not include instances of suboptimal drug therapy which may result in increased costs, extended length-of-stay, or ineffective therapy. Data in these areas are sparse, though medical care evaluations carried out as part of hospital quality assurance programs suggest that suboptimal therapy is common.

Other computer systems have been developed to influence physician decision making by monitoring patient data and providing feedback. However, most of these systems suffer from a significant structural shortcoming. This shortcoming involves the evaluation rules that are used to generate feedback. In all cases, these criteria consist of discrete, independent rules, yet medical decision making is a complex process in which many factors are interrelated. Thus, attempting to represent medical decision-making as a discrete set of independent rules, no matter how complex, is a task that can, at best, result in a first-order approximation of the process. This places an inherent limitation on the quality of feedback that can be provided. As a consequence it is extremely difficult to develop feedback that explicitly takes into account all information available on the patient. One might speculate that the lack of widespread acceptance of such systems may be due to the fact that their recommendations are often rejected by physicians. These systems must be made more valid if they are to enjoy widespread acceptance among physicians.

The MENTOR system is designed to address the significant problem of adverse drug reactions by means of a computer-based monitoring and feedback system to influence physician decision-making. It employs principles of artificial intelligence to create a more valid system for evaluating therapeutic decision-making.

The work in the MENTOR project is a collaboration between Dr. Blaschke at Stanford University, Dr. Speedie at the University of Maryland, and Dr. Charles Friedman at the University of North Carolina. Dr. Speedie provides the expertise in the area of artificial intelligence programming. Dr. Blaschke provides the medical expertise. Dr. Friedman contributes expertise in the area of physician feedback design and system impact evaluation. The blend of previous experience, medical knowledge, computer science knowledge and evaluation design expertise they represent is vital to the successful completion of the activities in the MENTOR project.

C. Highlights of Research Progress

The MENTOR project was initiated in December, 1983. The project has been funded by the National Center for Health Services Research since January 1, 1985. Initial effort focused on exploration of the problem of designing the MENTOR system. As of June 1, 1989, a working prototype system has been developed and is undergoing evaluation. The prototype consists of a Patient Data Base, an Inference Engine, an Advisory Module and a Medical Knowledge Base. The Medical Knowledge Base currently contains information related to aminoglycoside therapy, digoxin therapy, potassium supplementation, surgical prophylaxis, and microbiology lab reports. The system is currently implemented on a Xerox 1186 AI Workstation. Another version of the Patient Data Base has been developed for a VAXStation 3100 that is connected via an asynchronous line to the 1186 running the inference engine. The project has received additional funding from the National Center for Health Services Research to install and evaluate the MENTOR system in a Veterans Administration Hospital. This effort began in June of 1988 and will continue for two additional years. The VA system will reside on an 1186 and a VAX Station II connected directly to the VA's Ethernet LAN, and accessing hospital data through the FILEMAN software.

E. Funding Support

Title: MENTOR: Monitoring Drug Therapy for Hospitalized Patients

Principal Investigators:
Terrence F. Blaschke, M.D.
Division of Clinical Pharmacology
Department of Medicine
Stanford University

Stuart M. Speedie, Ph.D.
School of Pharmacy
University of Maryland

Funding Agency: National Center for Health Services Research
Grant Identification Number: 1 R18 HS05263
Total Award: January 1, 1985 - May 31, 1990 $1,091,750 Total
Direct Costs: Current Period: June 1, 1989 - May 31, 1990 $289,961 (Total Direct Costs)

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

This project represents a collaboration between faculty at Stanford University Medical Center, the University of Maryland School of Pharmacy, and the University of North Carolina in exploring computer-based monitoring of drug therapy. SUMEX, through its communications capabilities, facilitates this collaboration of geographically separated project participants by providing electronic mail and file exchange between sites.

B. Sharing and Interactions with Other SUMEX-AIM Projects

Interactions with other SUMEX-AIM projects has been on an informal basis. Personal contacts have been made with individuals working on the ONCOCIN project concerning system development issues. Dr. Perry Miller has also been of assistance by providing software for advisory generation. Given the geographic separation of the investigators, the ability to exchange mail and programs via the SUMEX system as well as communicate with other SUMEX-AIM projects is vital to the success of the project.

C. Critique of Resource Management

To date, the resources of SUMEX have been fully adequate for the needs of this project. The staff have been most helpful with any problems we have had and we are quite satisfied with the current resource management.

III. RESEARCH PLANS

A. Project Goals and Plans

The MENTOR project has the following goals:

1) Implement a prototype computer system to continuously monitor patient drug therapy in a hospital setting. This will be an expert system that will use a modular, frame-oriented form of medical knowledge, a separate inference engine for applying the knowledge to specific situations, and automated collection of data from hospital information systems to produce therapeutic advisories.
2) Select a small number of important and frequently occurring medical settings (e.g., combination therapy with cardiac glycosides and diuretics) that can lead to therapeutic misadventures, construct a comprehensive medical knowledge base necessary to detect these situations using the information typically found in a computerized hospital information system and generate timely advisories intended to alter behavior and avoid preventable drug reactions.
3) Design and begin to implement an evaluation of the impact of the prototype MENTOR system on physicians' therapeutic decision-making as well as on outcome measures related to patient health and costs of care.

1988 will be spent on continued prototype development in six content areas, refinement of the inference mechanisms, and installation of the system at the Palo Alto Veterans Administration Hospital.

B. Justification and Requirements for Continued SUMEX Use

This project needs continued use of the SUMEX facilities for one primary reason. Access to SUMEX is necessary to support the collaborative efforts of geographically separated development teams at Stanford and the University of Maryland.

Furthermore, the MENTOR project is predicated on the access to the SUMEX resource free of charge over the next two years. Given the current restrictions on funding, the scope of the project would have to be greatly reduced if there were charges for use of SUMEX.

C. Needs and Plans for Other Computing Resources Beyond SUMEX-AIM

A major long-range goal of the MENTOR project is to implement this system on a independent hardware system of suitable architecture. It is recognized that the full monitoring system will require a large patient data base as well as a sizeable medical knowledge base and must operate on a close to real-time basis. Ultimately, the SUMEX facilities will not be suitable for these applications. Thus, we have transported the prototype system to a dedicated